

## PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/199347>

Please be advised that this information was generated on 2019-06-02 and may be subject to change.

RESEARCH

Open Access



# Stable coevolutionary regimes for genetic parasites and their hosts: you must differ to coevolve

Faina Berezovskaya<sup>1</sup>, Georgy P. Karev<sup>2</sup>, Mikhail I. Katsnelson<sup>3</sup>, Yuri I. Wolf<sup>2</sup> and Eugene V. Koonin<sup>2\*</sup> 

## Abstract

**Background:** Genetic parasites are ubiquitous satellites of cellular life forms most of which host a variety of mobile genetic elements including transposons, plasmids and viruses. Theoretical considerations and computer simulations suggest that emergence of genetic parasites is intrinsic to evolving replicator systems.

**Results:** Using methods of bifurcation analysis, we investigated the stability of simple models of replicator-parasite coevolution in a well-mixed environment. We first analyze what appears to be the simplest imaginable system of this type, one in which the parasite evolves during the replication of the host genome through a minimal mutation that renders the genome of the emerging parasite incapable of producing the replicase but able to recognize and recruit it for its own replication. This model has only trivial or “semi-trivial”, parasite-free equilibria: an inefficient parasite is outcompeted by the host and dies off, whereas an efficient one pushes the host out of existence, leading to the collapse of the entire system. We show that stable host-parasite coevolution (a non-trivial equilibrium) is possible in a modified model where the parasite is qualitatively distinct from the host replicator in that the replication of the parasite depends solely on the availability of the host but not on the carrying capacity of the environment.

**Conclusions:** We analytically determine the conditions for stable coevolution of genetic parasites and their hosts coevolution in simple mathematical models. It is shown that the evolutionary dynamics of a parasite that initially evolves from the host through the loss of the ability to replicate autonomously must substantially differ from that of the host, for a stable host-parasite coevolution regime to be established.

## Reviewers

This manuscript was reviewed by Olivier Tenaillon, Sandor Pongor, and Alex Best. For the complete reports, go to the Reviewers' Reports section.

## Background

Genetic parasites are ubiquitous among cellular organisms [1]. In fact, most organisms host a variety of mobile genetic elements (MGE) that differ in their reproduction strategies and the mode of parasite-host interaction, including transposons, plasmids and viruses [2]. The abundance of the MGE in the biosphere is enormous. Viruses are by far the most common biological entities on earth [3–6], genes of MGE, such as those encoding transposases, are among the most abundant ones in diverse environments [7–9],

and the genomes of many multicellular organisms consist of up to 50% integrated MGE, in the case of mammals, or even up to 90% in the case of plants [10, 11].

The entire history of life can be properly depicted only as the perennial coevolution of cellular organisms with genetic parasites that includes both the proverbial arms race and various forms of cooperation [1, 12, 13]. Moreover, multiple lines of evidence point to a major role of genetic parasites in the evolution of biological complexity, in general, and in major transitions in evolution, in particular [14, 15].

The ubiquity and the enormous abundance of the MGE in the biosphere imply that genetic parasitism is an intrinsic feature of life. Indeed, parasites invariably emerge in computer simulations of the evolution of simple replicator systems which, in well-mixed models, typically leads to the collapse of the entire system [16–20]. This outcome can be avoided by incorporating compartmentalization into the model [18–20]. Further, mathematical models of

\* Correspondence: [koonin@ncbi.nlm.nih.gov](mailto:koonin@ncbi.nlm.nih.gov)

<sup>2</sup>National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD 20894, USA

Full list of author information is available at the end of the article



© The Author(s). 2018 **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated.

the evolution of genomes with integrated MGE, combined with probabilistic reconstruction of the evolution of bacterial and archaeal genomes, suggest that horizontal gene transfer at rates that are required to stave off the mutational meltdown of microbial populations (Muller's ratchet) prevents elimination of genetic parasites [21].

Cellular organisms and genetic parasites have been considered the two “empires” of life that fundamentally differ with regard to the capability of autonomous reproduction [22, 23]. The MGE fully depend on the host for energy production and the biosynthetic processes, in particular, translation (notwithstanding the fact that many large viruses encode components of the respective functional systems that modify and modulate the respective host functions [24, 25]).

In our previous theoretical analysis of the evolution of genetic parasites in simple replicator systems [26] which was, to a large extent, inspired by the seminal early experiments of Spiegelman and colleagues on reductive evolution of bacteriophages genomes in vitro [27–30], we present a semi-formal argument that the parasite-free state of a replicator is inherently evolutionarily unstable. Here, we explore these and derivative models analytically and show that, in order for a replicator and a parasite to stably coevolve, the parasite must substantially differ from the host in its reproduction strategy.

## Results and discussion

### Genetic parasite as a degraded variant of the replicator

A simple conceptual model of the emergence of a genetic parasite from within a self-replicating system [26] implies (initially infinitesimal) degradation of the replicase-encoding signal while the replicase-recognition signal is retained. The dynamics of such a system is described by the following pair of ordinary differential equations [26]:

$$\begin{aligned} \frac{dR}{dt} &= \frac{1}{(1+ae)} R^2 \left( 1 - \frac{R + \frac{P}{q}}{K} \right) - e_R R \\ \frac{dP}{dt} &= \frac{q}{(1+e)} RP \left( 1 - \frac{R + \frac{P}{q}}{K} \right) - e_P P \end{aligned} \quad (1)$$

(see Table 1 for the model parameters). The model is based on the following assumptions:

- The kinetics of replication depends on the interaction of the replicase and the template (either the host replicator or a parasite), so that the replication rate is proportional to the product of the respective population sizes ( $R^2$  for the replicator and

**Table 1** Parameters of the replicator-parasite dynamics model

Name	Description
$e_R$	Replicator intrinsic decay rate
$e_P$	Parasite intrinsic decay rate
$e$	Defense system efficiency
$a$	Cost of defense system
$q$	Parasite advantage factor
$K$	Carrying capacity of the environment

$RP$  for a parasite that lacks independent replication ability)

- The decay rates are constant ( $e_R$  and  $e_P$ ) for replicator and parasite respectively
- The parasite template replicates faster than that of the replicator by the factor  $q \geq 1$ , which, as the first approximation, could be interpreted as the “economy” factor (the simplest interpretation is parasite being  $q$  times smaller than the replicator, so its replication is faster by the factor of  $q$ )
- Both populations are environmentally limited by the same resources; the parasites consume  $q$  times less resources per individual compared to the host replicator; the environment carrying capacity  $K$  determines the point where replication becomes resource-limited
- The replicators possess a (costly) defense mechanism with efficiency  $e \geq 0$  that is capable to suppress the parasite replication by a factor of  $1 + e$  at the cost to the replicators replication rate of  $1 + ae$  (where  $a \geq 0$  is the defense mechanism cost factor)

As noticed previously [26], this system has equilibria ( $R = (K \pm \sqrt{K(K-4e_R)})/2$ ,  $P = 0$ ) as  $e = 0$ . The point  $R = (K - \sqrt{K(K-4e_R)})/2$ ,  $P = 0$  is always unstable, that is, introduction of the parasite into a replicator system without any defense near this equilibrium leads to eventual collapse of the entire system. The same happens at the equilibria  $R = (K + \sqrt{K(K-4e_R)})/2$ ,  $P = 0$  if  $q > e_P/e_R$ . If  $q < e_P/e_R$ , then this equilibrium is stable, and introduction of the parasite into the system leads to runaway parasite proliferation.

Here, we aim to identify all stationary states of model (1) and to analyze their stability, to study the qualitative behavior of this model and to suggest some modifications that make possible the stable host-parasite co-evolution.

All possible equilibria of the model can be found from the system of equations  $\frac{dP}{dt} = 0$ ,  $\frac{dR}{dt} = 0$ . These equations determine the following pair of (non-zero) isoclines (see Additional file 1: Mathematical Appendix 1 for details):

$$P_1(R) = \frac{q(R(K-R) - e_R K(1 + \alpha e))}{R} \quad (2)$$

$$P_2(R) = \frac{q R(K-R) - e_P K(1 + e)}{R}$$

the intersection of which ( $P_1(R) = P_2(R)$  at  $R > 0$ ,  $P > 0$ ) indicates a non-trivial equilibrium. It should be immediately apparent, however, that both equations have the same form and differ only by the values of the coefficients  $q(1 + \alpha e)e_R$  and  $(1 + e)e_P$ , respectively (see Fig. 1, left panel, for the characteristic shapes of these isoclines when  $q(1 + \alpha e)e_R \neq (1 + e)e_P$ ). Therefore, the curves do not intersect in any point ( $R > 0$ ,  $P > 0$ ); the only case where such equilibria exist is  $q = \frac{(1+e)e_P}{(1+\alpha e)e_R}$ ; in this case  $P_1(R) = P_2(R)$  over the whole range of  $R$  and the system has a line of non-isolated equilibria (see Fig. 1, right panel).

More formally, model (1) always has the trivial equilibrium  $O(R = 0, P = 0)$  (both populations collapse). In addition to that, the model might have up to two parasite-free, “semi-trivial” equilibria

$$O_1 \left( R = \frac{K - \sqrt{K(K - 4(1 + (1 + \alpha e)e_R))}}{2}, P = 0 \right) \quad (3)$$

$$O_2 \left( R = \frac{K + \sqrt{K(K - 4(1 + (1 + \alpha e)e_R))}}{2}, P = 0 \right)$$

the existence of which depends on the value of the following expression:

$$E = K - 4(1 + (1 + \alpha e)e_R) \quad (4)$$

If  $E > 0$ , both  $O_1$  and  $O_2$  exist; if  $E = 0$ ,  $O_1 = O_2 = O_{12}(R = K/2, P = 0)$ ; if  $E < 0$ , the model has only the trivial equilibrium  $O$ .

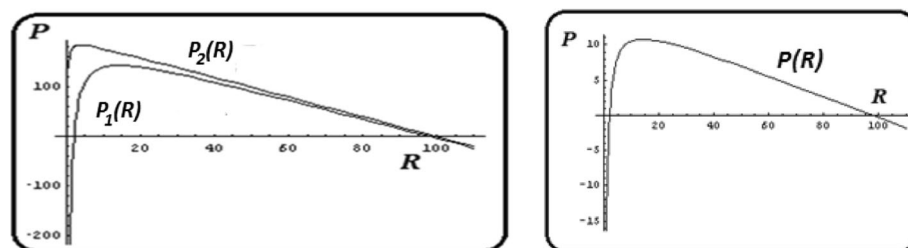
Stability analysis shows that the trivial equilibrium  $O$  is a stable node at all values of the model parameters. When both  $O_1$  and  $O_2$  exist ( $E > 0$ ), if  $q < \frac{(1+e)e_P}{(1+\alpha e)e_R}$ ,  $O_1$  is a saddle and  $O_2$  is a stable node and if  $q > \frac{(1+e)e_P}{(1+\alpha e)e_R}$ ,  $O_1$  is an unstable node and  $O_2$  is a saddle. If  $K = 4(1 + e)/q$ , the

only semi-trivial equilibrium  $O_{12}$  is a saddle-node fixed point (Fig. 2b; see Additional file 1: Mathematical Appendix 1 for details). Let us consider  $q$  and  $e$  as parameters of model (1) whereas  $K$ ,  $\alpha$ ,  $e_R$ ,  $e_P$  are arbitrary values of fixed coefficients. The parametric portrait of the system can be divided into 3 domains, **D1**, **D2** and **D3** with qualitatively different behaviors of the model (Fig. 2a); these domains are defined by Theorem 1 (Fig. 3 and Additional file 1: Mathematical Appendix 1). Asymptotically, as  $t \rightarrow \infty$ , there exist two qualitatively different types of behavior of the system, namely:

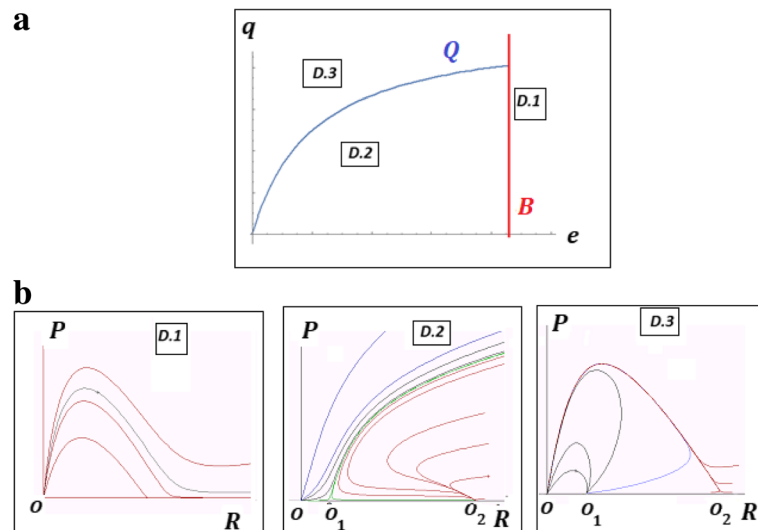
1. If the parameters  $q$ ,  $e$  belong to the domains **D1** or **D3** then, for any initial values ( $R, P$ ), the system collapses ( $R \rightarrow 0$ ,  $P \rightarrow 0$ );
2. If the parameters  $q$ ,  $e$  belong to the domain **D2**, then, there exist domains of initial values in which the system tends to the parasite-free equilibrium  $O_2$  whereas, at other initial values, it tends to the trivial equilibrium  $O$  (collapse).

Therefore, the original model [26] has no non-trivial ( $R > 0$ ,  $P > 0$ ) equilibria whereby the replicator and the parasite could coexist and coevolve. A stable parasite-free equilibrium can exist only if  $q < \frac{(1+e)e_P}{(1+\alpha e)e_R}$  and necessarily exists if, additionally,  $\alpha e < (K - 4e_R)/(4e_R)$ , that is, if the reproductive advantage of the parasite is low and the replicator possesses defense mechanisms that are efficient enough and not too costly to overcome the intrinsic advantage of the parasite.

The implications of these findings are the following: emergence of a genetic parasite as a (slightly) degraded copy (e.g. a variant of an RNA genome that contains a deletion and thus cannot produce an active replicase) of the “naïve” (defenseless) replicator (i.e. when eqs. (1) apply and  $e = 0$ ) leads to the collapse of the system towards the trivial equilibrium ( $R = 0$ ,  $P = 0$ ). If the replicator already has a sufficiently advanced and not



**Fig. 1** Non-trivial null-clines  $P_1(R)$ ,  $P_2(R)$  for model (1). See eq.(2). Left panel: for  $q(1 + \alpha e)e_R \neq (1 + e)e_P$ , the isoclines  $P_1(R)$ ,  $P_2(R)$  do not intersect for positive  $P$ ,  $R$ . Right panel:  $P_1(R) \equiv P_2(R)$  for  $q(1 + \alpha e)e_R = (1 + e)e_P$ , so system (1) has a line of non-isolated equilibria



**Fig. 2** Characteristics of model (1). **a** Parameter-phase portrait. **D1**, **D2** and **D3** are the domains with qualitatively different model behaviors. The boundaries are: **Q**:  $q = \frac{(1+e)e_p}{(1+ae)e_R}$  (blue), **B**:  $e = \frac{K-4e_R}{4ae_R}$  (red). **b** Examples of phase portraits of model (1) with a stable trivial equilibrium  $O$  (all domains); stable semi-trivial equilibrium  $O_2$  and unstable semi-trivial equilibrium  $O_1$  (domain **D2**); unstable semi-trivial equilibria  $O_1$ ,  $O_2$  (domain **D3**)

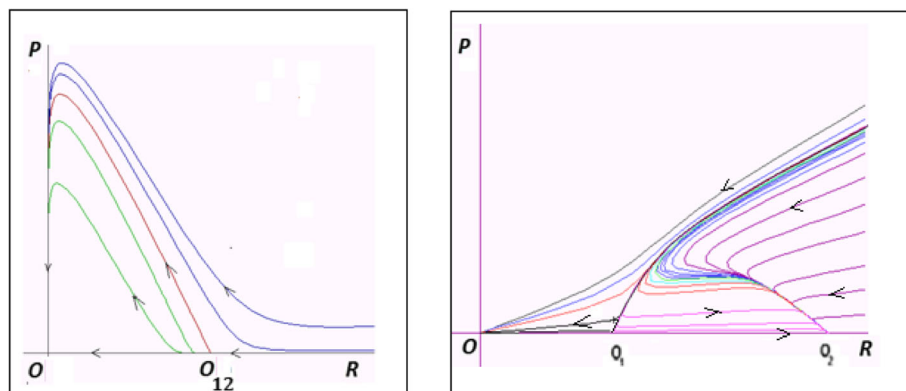
excessively costly defense mechanism (or is able to evolve it before the population collapse), the system could be stable in the parasite-free state as long as the defense mechanism persists. If the defense degrades over time, which is likely to be the case because most if not all defense mechanisms incur a non-zero fitness cost [31–34], the system will become vulnerable again. Therefore, such a system is inherently unstable. If the early history of (pre-)life included a primitive replicator stage, as the RNA World concept implies [17, 35, 36], it would be vulnerable to parasite-driven collapse, could not have persisted for a long time and necessarily would evolve into a different mode of replicator-parasite relationships that is not subject to the limitations imposed by eq. (1).

#### Genetic parasite with an additional interaction with the replicator

Let us consider the model that differs from the model (1) by an additional effect of the replicator-parasite interaction on the replicator dynamics:

$$\begin{aligned} \frac{dR}{dt} &= \frac{1}{(1+ae)} R^2 \left( 1 - \frac{R + \frac{P}{q}}{K} \right) - bRP - e_R R \\ \frac{dP}{dt} &= \frac{q}{(1+e)} RP \left( 1 - \frac{R + \frac{P}{q}}{K} \right) - e_P P \end{aligned} \quad (5)$$

Such an effect could be interpreted as an extra penalty on the replication rate:



**Fig. 3** Phase portraits of the model (1) on the boundaries **B** (left panel) and **Q** (right panel) of the phase-parametric portrait

$$\left[ R \frac{1}{(1+ae)} \left( 1 - \frac{R + \frac{P}{q}}{K} \right) - bP \right] R$$

and/or as extra replicator degradation rate:

$$[e_R + bP]R$$

due to the replicator-parasite interaction.

The trivial equilibrium  $O$  ( $R=0$ ,  $P=0$ ), and the semi-trivial equilibria  $O_1$  and  $O_2$  ( $R>0$ ,  $P=0$ ), identified for the model (1) still exist in the system (5) under the same conditions. In addition, however, two new equilibria also can exist in the system, namely,  $A_1(R^+, P^*)$  and  $A_2(R^-, P^*)$  where

$$R^{+,-} = \frac{q(1+ae)e_R - (1+e)e_P + bKq^2(1+ae)}{\pm \sqrt{((1+e)e_P - (1+ae)q(e_R + bKq))^2 - 4(1+e)b^2(1+ae)^2 e_P K q^3}} \\ 2b(1+ae)q^2$$

$$P^* = \frac{(1+e)e_P - q(1+ae)e_R}{bq(1+ae)}$$

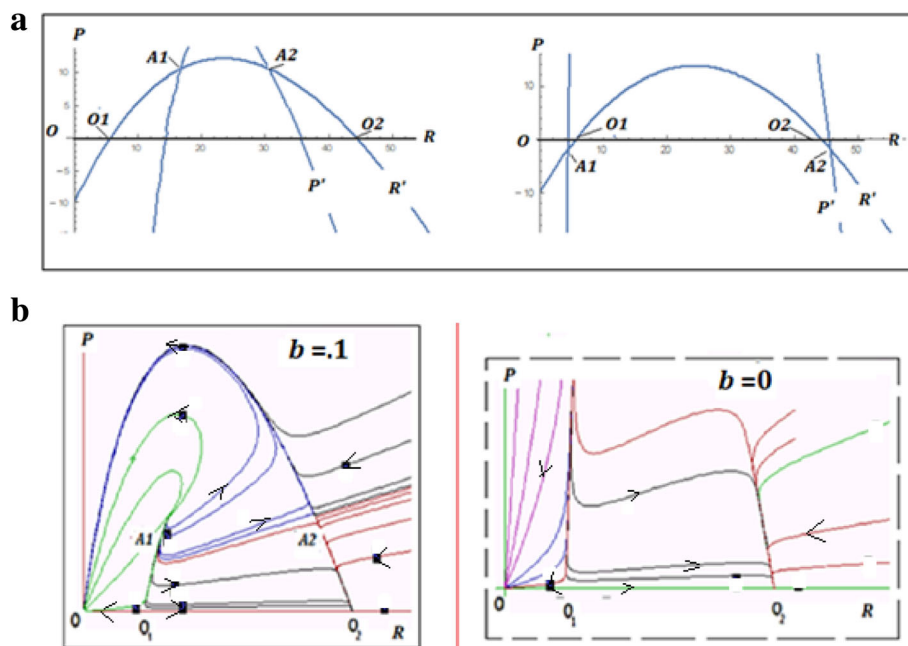
These equilibria are positive (i.e. imply  $P>0$ ) only if  $q < \frac{(1+e)e_P}{(1+ae)e_R}$ ; if  $q = \frac{(1+e)e_P}{(1+ae)e_R}$ , the equilibrium points  $A_1$  and  $A_2$  merge with the points  $O_1$  and  $O_2$  of eq. (3) and become negative if  $q > \frac{(1+e)e_P}{(1+ae)e_R}$ .

Stability analysis, however, shows that, when these equilibria belong to the first quadrant, both are unstable:  $A_1$  is an *unstable node* and  $A_2$  is a *saddle* (Fig. 4a, b; see

Additional file 2: Mathematical Appendix 2 for details). Therefore, although the inclusion of this extra interaction, corresponding to a more aggressive parasite compared to model (1), into the model does lead to non-trivial equilibria (i.e. replicator-parasite coexistence), they are unlikely to persist for on the evolutionarily relevant time scale.

### A highly derived genetic parasite

Analysis of the model (1) shows that the equilibrium between the parasite and the replicator is unstable and largely hinges on the relative growth advantage of the parasite  $q$  and the efficiency of the host replicator's defense mechanisms  $\frac{(1+e)}{(1+ae)}$ . A parasite that has a greater advantage over the host replicator than the defense mechanisms can handle overwhelms the system and drives it to collapse, whereas a less efficient parasite is eliminated by the host. Simply shrinking the parasite genome and hence increasing its replicative advantage ( $q \gg 1$ ), while concomitantly reducing the impact on the replicator dynamics according to the expression  $1 - \frac{R + \frac{P}{q}}{K}$ , does not represent a viable path for the parasite towards the stable coexistence because the increase in the parasite replication efficiency overwhelms the attenuation of its deleterious effect. Moreover, as the analysis of the model (5) indicates, another intuitively plausible path to coexistence, through additional suppression of the replicator (allowing the parasite to escape elimination in the



**Fig. 4** Characteristics of model (5). **a** Null - isoclines and equilibria. Left panel: positive non-trivial equilibria  $A_1$ ,  $A_2$  as  $q=4$ ; right panel: no positive non-trivial equilibria as  $q=10$ . **b** Phase portraits. Left panel: phase portrait of model (5) as  $b>0$  with unstable non-trivial equilibria  $A_1$  and  $A_2$ . Right panel: phase portrait of model (5) as  $b=0$  is similar to the phase portrait of model (1), no non-trivial equilibria



presence of the efficient defense mechanism, for example, by evolving antidefense mechanisms) does not work either. Although non-trivial equilibria can exist in this model, they are unstable.

More generally, the equilibria appear in the system as intersections of the isoclines (see eq. (2)). Thus, to ensure stable coexistence of the host and the parasite, the equations of the replicator and parasite dynamics should substantially differ from each other unlike those in model (1) (see Additional file 1: Mathematical Appendix 1). From the biological perspective, this means that the parasite cannot be, simply, a slightly modified variant of the replicator. Rather, the effects of the replicator and the parasite on each other's replication should be substantially asymmetric and/or their interactions with the environment should be substantially different. One such possible modification is represented in the following model:

$$\frac{dR}{dt} = \frac{1}{(1+ae)} R^2 \left( 1 - \frac{R + \frac{P}{q}}{K} \right) - e_R R \quad (6)$$

$$\frac{dP}{dt} = \frac{q}{(1+e)} R P - e_P P$$

In the model (6), the parasite dynamics does not depend on the carrying capacity of the environment, whereas the consumption of the resources by the parasite continues to be a factor in the replicator dynamics ( $1 - \frac{R + \frac{P}{q}}{K}$ ). A biological model for this behavior is a highly specialized parasite that obtains the necessary resources from the host (effectively, for free) rather than directly from the environment. If the parasite is individually small enough compared to the host, as is the case for many viruses and transposons, at least those that parasitize on eukaryotes, its growth is not limited by the external environmental resources. The mere availability of the host ensures that the resources are sufficiently abundant and is the only limiting factor for the parasite reproduction.

Analysis of the model equilibria (see Additional file 3: Mathematical Appendix 3 for details) defines the following family of null isoclines for  $P$  and  $R$ :

$$\begin{aligned} P_1 &= 0 \\ P_2(R) &= \frac{q R(K-R) - e_P K(1+e)}{R} \\ R_1 &= 0 \\ R_2 &= \frac{(1+e)e_P}{q} \end{aligned} \quad (7)$$

(Fig. 5). The trivial equilibrium  $O$  ( $R=0$ ,  $P=0$ ), and the semi-trivial equilibria  $O_1$  and  $O_2$  ( $R>0$ ,  $P=0$ ), found in model (1) and defined by eq. (3), also exist in model

(6) subject to the conditions defined by eq. (4). In addition, an equilibrium

$$A(R_A = \frac{(1+e)e_P}{q}, P_A = \frac{Kq(1+e)e_P - (1+e)^2 e_P^2 - (1+ae)e_R K q^2}{(1+e)e_P}) \quad (8)$$

also exists. Under the condition  $q = \frac{(1+e)e_P(K + \sqrt{K(K-4(1+ae)e_R})})}{2(1+ae)e_R K}$ , the equilibrium  $A$  is semi-trivial ( $R>0$ ,  $P=0$ ) and coincides with either  $O_1$  or  $O_2$ . For the values of  $q$  defined by

$$\begin{aligned} &\frac{(1+e)e_P(K - \sqrt{K(K-4(1+ae)e_R})})}{2(1+ae)e_R K} < q \\ &< \frac{(1+e)e_P(K + \sqrt{K(K-4(1+ae)e_R})})}{2(1+ae)e_R K} \end{aligned}$$

the equilibrium  $A$  is non-trivial ( $R>0$ ,  $P>0$ ).

Analysis of the model (6) (see Additional file 3: Mathematical Appendix 3 for details) allows us to construct the parameter-phase portrait of the model (Fig. 6a, b). Let us consider  $q$  and  $e$  as parameters of the model (6) whereas  $K$ ,  $\alpha$ ,  $e_R$ ,  $e_P$  are arbitrary values of fixed coefficients. The portrait contains domains **D1**, **D2** and **D3** similar to those in model (1) (Fig. 2), and 3 additional domains, **D4**, **D5** and **D6**.

The boundaries between the parametric domains of model (6) are determined by the curves  $B, Q_1, Q_2, Tr, L$ , described in Additional file 3: Mathematical Appendix 3.

The domains have the following properties:

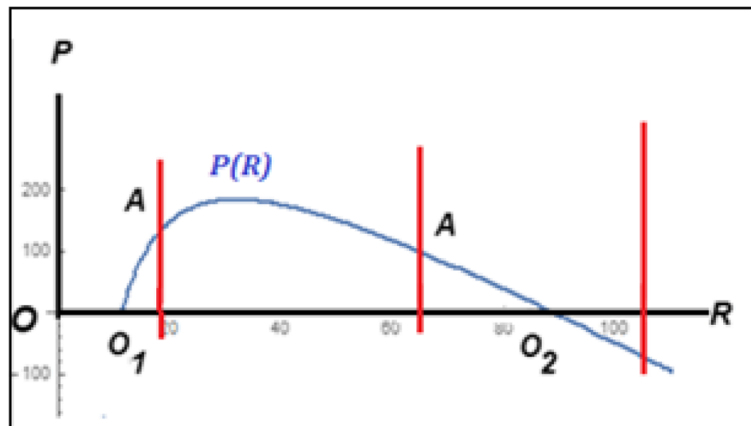
**D1** ( $e > B$ ): only the trivial equilibrium  $O$  exists and is stable (in **D1**, the effective cost of the defense mechanisms  $ae$  is so high that the replicator population collapses even in the absence of the parasite);

**D2** ( $e < B$ ,  $q < Q_1$ ): the trivial equilibrium  $O$  is stable; the semi-trivial equilibrium  $O_1$  is unstable; the semi-trivial equilibrium  $O_2$  is stable; no other equilibria exist (in **D2**, the efficiency of the defense mechanisms is sufficient to eliminate the parasite as long as the population of the replicator itself is sustainable);

**D3** ( $e < B$ ,  $q > Q_2$ ): the trivial equilibrium  $O$  is stable; both semi-trivial equilibria  $O_1$  and  $O_2$  are unstable; no other equilibria exist (in **D3**, the parasite overwhelms the defense mechanisms of the replicator and thus collapses the system);

**D4** ( $e < B$ ,  $Q_1 < q < Tr$ ): the trivial equilibrium  $O$  is stable; both semi-trivial equilibria  $O_1$  and  $O_2$  are unstable; the non-trivial equilibrium  $A$  is a stable node or focus (in **D4**, the parasite and the replicator can coexist at a stable equilibrium, Fig. 6b);

**D5** ( $e < B$ ,  $Tr < q < \min(L, Q_2)$ ): the trivial equilibrium  $O$  is stable; both semi-trivial equilibria  $O_1$  and  $O_2$  are unstable; the non-trivial equilibrium



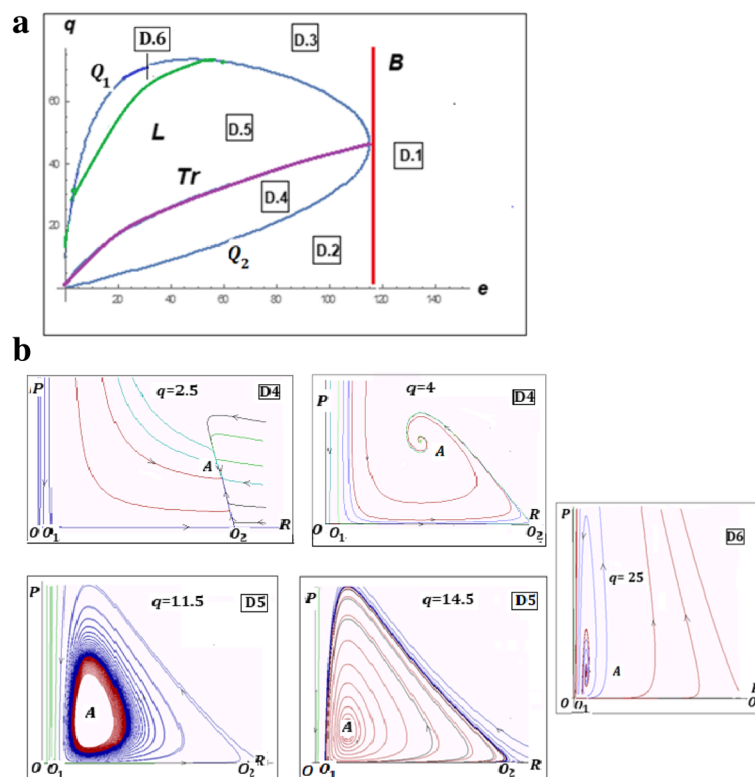
**Fig. 5** Null isoclines (eq. (7)) for model (6)

$A$  is unstable but is surrounded by a stable limit cycle (in  $D5$ , the parasite and the replicator can coexist in a stable oscillation regime);

$D6$  ( $e < B$ ,  $L < q < Q_2$ ): the trivial equilibrium  $O$  is stable; both semi-trivial equilibria  $O_1$  and  $O_2$  and the non-trivial equilibrium  $A$  are unstable (in  $D6$ , the defense mechanisms of the replicator are not sufficient to keep the parasite in check; although other equilibria

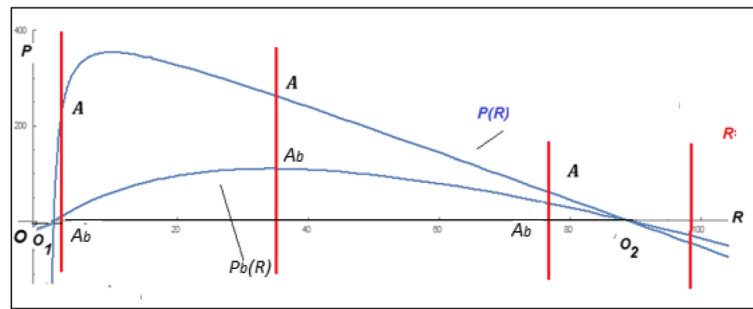
exist, they are unstable, so that the system would collapse upon a perturbation).

Overall, the system collapses in domains  $D1$ ,  $D3$ , and  $D6$ ; the replicators and parasites can coexist in domains  $D4$  and  $D5$ ; and, the system is bistable in domains  $D2$ ,  $D4$  and  $D5$ , that is, its final behavior critically depends on the initial values of  $P$  and  $R$ .



**Fig. 6** Characteristics of model (6). **a** Phase-parametric portrait of model (6).  $D1 - D6$  are the domains with qualitatively different model behaviors. **b** Examples of phase portraits of model (6) with a stable non-trivial equilibrium  $A$  (domain  $D4$ ) and stable oscillations (domain  $D5$ ); no non-trivial stable regimes in domain  $D6$





**Fig. 7** Null isoclines for Volterra-type models. The isocline  $P(R)$  corresponds to the model (A3.1) which is the same as model (6); the isocline  $P_b(R)$  corresponds to the model (A4.1) with  $b > 0$

Thus, under the model (6), there exists an intermediate regime in which the parasite efficiency is roughly balanced by the efficiency of the replicator defense mechanisms under which the replicator and the parasite can coexist indefinitely in a stable equilibrium or a limit cycle. Unlike in the model (1) that lacks such a regime, the parasite in the model (6) is a highly derived state. In such a state, the parasite replication and survival directly depend only on the host (replicator) availability but not on the environment (the parasite equation in (6) lacks the  $1 - \frac{R+p}{K}$  term). Models with the parasite being closely similar to the replicator in its interaction with the environment (that is, with the parasite replication equation closely resembling that of the replicator), such as the model (1), cannot support stable non-trivial equilibria (compare the families of null isoclines in eqs. (2) and (7)). A variant of the model (6) with an additional effect of the replicator-parasite interaction on the replicator dynamics shows qualitatively the same behavior as model (6) (see Fig. 7 and Additional file 4: Mathematical Appendix 4).

### Concluding remarks

We show here that, in order to produce stable equilibria, models of the coevolution of genetic parasites and their hosts cannot be too simple. The parasites must not be too closely similar to the host in its reproduction strategy. More specifically, stable coevolution becomes possible in models where only the reproduction of the host but not that of the parasite depends on the carrying capacity of the environment. From a biological perspective, a successful parasite has to rely on the host not only for replication but also for building blocks and energy. Perhaps, these results go some way to explain why, to the best of our current knowledge, no genetic parasites have ever captured neither full-fledged biosynthetic pathways, including the translation system, nor the molecular machinery for energy production.

Although the analyzed models are too simple to be of much direct relevance to extant cases of host-parasite

coevolution, they are likely to be relevant for early stages of (pre) life evolution, within the RNA World, the leading current scenario for the origin of life [37–39], and at the subsequent stages, when DNA and proteins came to the scene, and different replication strategies evolved. Importantly, the replicators in this model are not abstract information carriers but self-sustaining reproducers that extract energy and building blocks from the environment, supporting both the host and the parasite [17, 35]. Conversely, the parasites neither encode their own replication machinery nor actively utilize resources, that is, their dependence on the host is complete.

The results of this work imply that primordial replicators have made innumerable “false starts” whereby host-parasite systems collapsed under the unchecked parasite pressure. Only in more evolved systems, where the capacity of the parasite to outcompete the host is balanced by defense mechanisms and, conversely, the ability of the host to eliminate the parasites is undermined by the cost of defense, stable coevolution became possible. Such coevolution between hosts and parasites is likely to be an essential driver of the evolution of biological complexity, and more specifically, of major transitions in evolution [15]. Hence any biological systems, in which a stable host-parasite coevolution regime failed to evolve, would be evolutionary dead ends.

We considered here only homogenous, well-mixed host-parasite systems. In computer simulations, compartmentalization has been shown to lead to stable co-evolutionary regimes. Most likely, compartmentalization of replicator ensembles had been part and parcel of the evolution of life from its earliest stages on, and could be considered a form of host defense, perhaps, the simplest one [18, 19]. Clearly, compartmentalization is a major path to the emergence of diversity and complexity. Further development of the models described here, in particular, by explicitly allowing evolution of the parameters defining the behavior of the parasites and replicators, will be of interest.

## Reviewers' reports

### Reviewer 1: Olivier Tenaillon

In their manuscript “the Stability of host parasite systems: you must differ to coevolve” Faina Berezovskaya and coauthors study the existence of non trivial equilibrium in a host parasite system. The model they are interested in is the emergence of virus, transposons or mobile genetic elements. Using the existing sets of equation, the model does not allow for coexistence but rather suggests only non-invasion or the systems collapse due to parasite. TO find a range of conditions in which host and parasite can coexist they show that allowing the parasite to fully capture the energy from its host is important. This suggest that parasite must from the start be very different from their host. The problem is nicely stated and important for the current understanding of genomes that are full of mobile genetic elements. I have just some minor comments concerning how the equations could be reduced at least in the end to clearly identify the relevant biological parameters.

First, in the set of equation used  $K$  is not the capacity at equilibrium, would it be possible to have a true carrying capacity term in the equation for  $R$  and add modification in the  $P$  equation.

Response: *The terms for the environment interactions for the host and the parasite were deliberately entangled in the original formulation of the model (eq. (1), ref [26]). In this model, the host and the parasite compete for the same resources (albeit consume different amounts per individual, see more on this below). That is where the  $\frac{R+P/q}{K}$  term comes from. The carrying capacity  $K$ , indeed, is not equal to the total population size at equilibrium. It defines the neutral point for the resource-limited replication rate rather than for the overall population growth rate. We find this usage to be quite in the spirit of the original logistic model; the alternative (defining  $K$  through the total population size at equilibrium) would entangle it with the decay term and make the equations more cumbersome. We have added the explanation to this effect to the main text.*

The impact of the  $q$  term in the logistic is not clear to me. It should be explained why  $q$  increases the speed of the parasite and decreases its contribution in the logistic. Though it makes sense qualitatively, could decoupling  $q$  in two terms solve some of the problem? Here one of the issue is that a small impact on the host resources, results in an infinite growth rate. There could be a decoupling of the two with a replication rate limit for instance, a simple saturating function could work.

Response: *The simplest way to think of the term  $q$  is to envisage a primitive system where both the replicator and the parasite are, essentially, single molecules as in the RNA world scenario. The replicator molecule encodes the*

*complete, active replicase, whereas the parasite molecule only carries the replication recognition signals and is shorter by a factor of  $q$ . Replication of such a parasite molecule requires  $q$  times less resources and would occur  $q$  times faster, compared to the replicator molecule. We have added the explanation to this effect to the main text.*

In the final set of equation analysed (eqs. 6), here again it would be interesting to narrow down the system to its simplest form, to be able to see how general these equations are and how important are the terms linked to resistance and cost. For instance, the Parasite equation results in a pure predator equation. But the impact on the host/prey is different, mostly due to the  $R^2$  in the replicator eq. I think discussing the similarity and reducing the final numbers of parameters could be useful for readers and for the discussion.

Response: *The  $R^2$  term in the replicator equation and the  $RP$  term in the parasite equation both come from a simple mechanistic consideration: the entity (either the host replicator or a parasite) must encounter a replicase in order to be replicated. This is what makes this system distinct from the classic predator-prey model where the prey population growth follows the first-order kinetics (i.e. is not limited by the rate of intra-species encounters). We consider this feature to be an important distinction. We have added the explanation to this effect to the main text.*

A lot is spent about discussing the resistance to the parasite, but the final equation could just assume different growth rates. For the discussion, in the debate on parasite emergence, here a choice is clearly made in the equations: the parasite has to be fully dependent on its hosts from the start. What if as observed in some viral system [40] there is only an initial partial dependency. This could lead to an unstable system with a prisoner's dilemma outcome. But with a group selection pattern, the gradual emergence of parasites that can coexist with their host could emerge. For the present manuscript I just think that it could be mentioned that the present systems envision the immediate appearance of a fully dependent parasite, but that more gradual outcomes could be possible. Overall it is an interesting story.

Response: *We added the clarification that the model presumes a complete dependence of the parasite on the replicator in the model description and also in the Concluding Remarks.*

### Reviewer 2: Sandor Pongor

This is an original and significant contribution. As pointed out to the authors, I find the language difficult at times, though the writing of the paper is in general excellent and easy to follow.

Genetic parasites including mobile genetic elements, viruses and plasmids are ubiquitous among cellular

organisms, and are by far the most abundant form of life. The emergence of genetic parasites is intrinsic to evolving replicator systems, but the stability of host-genetic parasite coevolution are not well understood. Berezovskaya and associates tackle this very complex problem using the method of bifurcation analysis. They convincingly show that the logically simplest model is unstable however stability can be achieved via slight modification of this model. The logics of the presentation is straightforward and easy to follow.

This reviewer finds the language too complicated at times, for instance the key sentence of the abstract “It is shown that the simplest imaginable system of this type ...” appears far too complex for and abstract, and the subtitle “A highly derived genetic parasite” is hard to follow.

Response: *We appreciate the review and regret that some of the original language appeared difficult in places. We streamlined the complicated sentence in the Abstract and went through the entire manuscript to split some long sentences which, hopefully, makes the text more digestible. We presume that “subtitle” here implies the text of the respective section (the subtitle itself is very brief and straightforward). We made a few edits in this section but, overall, fail to see any source of potential confusion.*

### Reviewer 3: Alex Best

The authors present a series of models to represent dynamics between a host replicator and a genetic parasite. They show that their initial models yield no stable coexistence equilibria, but that if the parasite derives its resources purely from the host, and not the environment, then stable or cycling coexistence can occur. My background is in host-parasite coevolution in the more ‘classic’ setting of microparasites, using S-I type epidemiological models (after Kermack & McKendrick and Anderson & May). I was therefore very interested by the idea of looking at the case of genetic parasites and MGEs. The methods applied look sound and the manuscript was fairly easy to read. Overall, however, I find myself a little confused. The set-up of the ‘simplest imaginable’ model seems to me rather less an obvious starting point than their final model. This perhaps represents my background in the microparasite literature, but even so I think more could be done to explain \*why\* this would be an obvious starting point for a model.

The assumption of equal carrying capacities and why replicator growth is proportional to  $R^2$  is particularly puzzling to me. I also found the figures to be of quite low print-quality, and therefore not always very helpful.

Response: *The carrying capacity in the original model is determined by the joint resource consumption whereby the replicator consumes 1 unit per individual and the parasite  $1/q$  units per individual. In our model, the host*

*replicator and the host consume the same resource and hence share a common carrying capacity although the resource consumption rates differ by a factor of  $q$ . Such a model would apply, for example, to a case when the replicator and the parasite are both molecules of the same type, such as RNA. The  $R^2$  dynamics for the replicator and the RP dynamics for the parasite reflects the second order kinetics where the act of replication requires an encounter between a template and the replicase. We added more detailed explanations to the model description.*

Title - Could you perhaps use ‘genetic parasites’ in the title? This is a fairly specific example of a ‘host-parasite system’. Figures - The figure legends are far too brief, and the plots themselves seem to be of low quality.

Response: *Yes, fair point, we have modified the title accordingly and also emphasized genetic parasites in the revised Conclusions section of the Abstract. Several figure legends have been expanded, too.*

P5 L33 - These equations do not seem to be the obvious first step, and I’m not sure how they were reached. For example: \* While it makes some sense for the parasite growth rate to be proportional to the replicator density, why is the replicator growth rate proportional to  $R^2$ ? This seems particularly important as it will be this assumption that ensures the (0,0) equilibrium is always stable and this complete extinction remains possible.

\* Why would we expect the parasite to be subject to the same carrying capacity,  $K$ , as the replicator?

Response: *See the explanation above.*

P5 L41 - Table 1 doesn’t seem to be here anywhere.

Response: *we regret the omission, table added.*

P5 L43 - What is ‘the template’? You haven’t defined this previously.

Response: *We use the word “template” to mean “anything that could be replicated by the replicator”. In the context of the model, it is either the replicator itself that is replicated by a replicase (which, in a simplest model of the RNA world, could be an identical molecule), or a parasite (which is incapable of producing the replicase but can serve as a template). In molecular biology, “template” is a common term when replication (or transcription) is described, so we do not expect any confusion coming out of it.*

P6 L10 -  $R = K$  is not an equilibrium of (1), as this would give  $dR/dt = -e_r K$ . I assume that is why you use the  $\approx$  symbol, but why not state the actual value.

Response: *The statement is corrected to refer to the equilibrium of a parasite-free system.*

P6 L53 - Why do you call these ‘semi-trivial’?

Response: *Because the value of one of the variables (but not both) is 0. When introducing the semi-trivial equilibrium, we indicate that it is parasite-free, so we do not believe any confusion is likely.*

P7 L31 - You have not defined these domains.

Response: *they are defined in the revised text, with reference to Additional file 1: Mathematical Appendix 1 and Fig. 2a.*

P7 L58 - In what sense is the parasite a ‘slightly degraded copy’ of the replicator?

Response: *we specify in the revision: “e.g. a variant of an RNA genome that contains a deletion and thus cannot produce an active replicase”.*

P9 L18 - In Fig. 4, these new coexistence equilibria seem to introduce an interesting bistability between replicator-phy and complete extinction. Could you comment on this?

Response: *Positive non-trivial equilibria A1(left) and A2 (right) (see Fig. 4) that appear when  $q < \frac{(1+e)e_p}{(1+ae)e_R}$  are both unstable: A1 is an unstable node and A2 is a saddle. Stable equilibria are only trivial (O) and semi-trivial (O2) - see Fig. 2, Domain 2; new stable equilibria do not appear. Unstable equilibria are not “observable” but can change the areas of attraction of stable equilibria (in particular, the area of attraction of the semi-trivial equilibrium O2 increases) and make the system orbits more complex.*

P10 L17 - As I implied above, it is a little puzzling to me why this wouldn’t be the first model you choose anyhow.

Response: *We started from the model (1) in order to show that, if the dynamics of genetic parasites is close to that of the host, then, either the parasite or the entire system goes extinct.*

Why would you assume the replicator and parasite have the same carrying capacity? This model is much closer to ‘classic’ host-parasite models.

Response: *it is not exactly the case that the carrying capacities for the replicator and the parasite are the same. Rather, K is the carrying capacity for the complete host-parasite system from which they consume resources at different rates (for more detail, see the response to Olivier Tenaillon above).*

P12 L10 - It is very interesting that you have found a limit cycle here, which does not occur in classic models. My guess is this is again to do with the  $R^2$  in the growth term, but it would be interesting if you could explain this.

Response: *A limit cycle arises due to the term  $R^2$  in the equation for the replicator growth that differentiates the proposed model from standard host – parasite models. As a result, the relative growth rate  $(dR/dt)/R$  is non-monotonic and has a maximum, so that the cycle arises when the null-cline intersects this maximum. Similar models describe so-called Allee-effect [41]. In this type of models, oscillations always arise for such parameter values where a stable equilibrium loses stability and a stable limit cycle appears around the equilibrium due to the Hopf bifurcation. This limit cycle can disappear in a*

*heteroclinics composed by separatrices of neighboring saddles under variation of the system parameters, and this is exactly what happened in our model (6). As a result, only unstable non-trivial equilibrium remains in the phase plane. Then, with further changing of the parameters, this equilibrium merges with the trivial equilibrium and disappears.*

## Additional files

**Additional file 1:** Mathematical Appendix 1. (DOCX 25 kb)

**Additional file 2:** Mathematical Appendix 2. (DOCX 16 kb)

**Additional file 3:** Mathematical Appendix 3. (DOCX 28 kb)

**Additional file 4:** Mathematical Appendix 4. (DOCX 19 kb)

## Abbreviation

MGE: Mobile genetic elements

## Funding

The authors’ research is supported by intramural funds of the US Department of Health and Human Services (to the National Library of Medicine).

## Availability of data and materials

Supporting data are available as additional files.

## Authors’ contributions

FB and GPK performed research; FB, GPK, YIW, MIK, and EVK analyzed the results; GPK, YIW, and EVK wrote the manuscript that was edited and approved by all authors.

## Ethics approval and consent to participate

Not applicable.

## Consent for publication

Not applicable.

## Competing interests

The authors declare that they have no competing interests.

## Publisher’s Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Author details

<sup>1</sup>Department of Mathematics, Howard University, Washington, DC 20059, USA. <sup>2</sup>National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD 20894, USA. <sup>3</sup>Institute for Molecules and Materials, Radboud University, 6525AJ Nijmegen, Netherlands.

Received: 27 August 2018 Accepted: 29 November 2018

Published online: 14 December 2018

## References

- Koonin EV, Dolja W. A virocentric perspective on the evolution of life. *Curr Opin Virol*. 2013;3(5):546–57.
- Koonin EV, Starokadomskyy P. Are viruses alive? The replicator paradigm sheds a decisive light on an old but misguided question. *Stud Hist Phil Biol Biomed Sci*. 2016; in press.
- Rohwer F. Global phage diversity. *Cell*. 2003;113(2):141.
- Rohwer F, Thurber RV. Viruses manipulate the marine environment. *Nature*. 2009;459(7244):207–12.
- Suttle CA. Viruses in the sea. *Nature*. 2005;437(7057):356–61.
- Suttle CA. Marine viruses—major players in the global ecosystem. *Nat Rev Microbiol*. 2007;5(10):801–12.



7. Dahlberg C, Hermansson M. Abundance of Tn3, Tn21, and Tn501 transposase (tnpA) sequences in bacterial community DNA from marine environments. *Appl Environ Microbiol*. 1995;61(8):3051–6.
8. Brazelton WJ, Baross JA. Abundant transposases encoded by the metagenome of a hydrothermal chimney biofilm. *ISME J*. 2009;3(12):1420–4.
9. Vigil-Stenman T, Ininbergs K, Bergman B, Ekman M. High abundance and expression of transposases in bacteria from the Baltic Sea. *ISME J*. 2017; 11(11):2611–23.
10. Gogvadze E, Buzdin A. Retroelements and their impact on genome evolution and functioning. *Cell Mol Life Sci*. 2009;66(23):3727–42.
11. Solyom S, Kazazian HH Jr. Mobile elements in the human genome: implications for disease. *Genome Med*. 2012;4(2):12.
12. Forterre P, Prangishvili D. The great billion-year war between ribosome- and capsid-encoding organisms (cells and viruses) as the major source of evolutionary novelties. *Ann N Y Acad Sci*. 2009;1178:65–77.
13. Forterre P, Prangishvili D. The major role of viruses in cellular evolution: facts and hypotheses. *Curr Opin Virol*. 2013;3(5):558–65.
14. Iranzo J, Lobkovsky AE, Wolf YI, Koonin EV. Virus-host arms race at the joint origin of multicellularity and programmed cell death. *Cell Cycle*. 2014;13(19): 3083–8.
15. Koonin EV. Viruses and mobile elements as drivers of evolutionary transitions. *Philos Trans R Soc Lond Ser B Biol Sci*. 2016; in press.
16. Szathmari E, Demeter L. Group selection of early replicators and the origin of life. *J Theor Biol*. 1987;128(4):463–86.
17. Szathmari E, Maynard Smith J. From replicators to reproducers: the first major transitions leading to life. *J Theor Biol*. 1997;187(4):555–71.
18. Takeuchi N, Hogeweg P. The role of complex formation and deleterious mutations for the stability of RNA-like replicator systems. *J Mol Evol*. 2007; 65(6):668–86.
19. Takeuchi N, Hogeweg P. Evolutionary dynamics of RNA-like replicator systems: a bioinformatic approach to the origin of life. *Phys Life Rev*. 2012;9(3):219–63.
20. Takeuchi N, Hogeweg P, Koonin EV. On the origin of DNA genomes: evolution of the division of labor between template and catalyst in model replicator systems. *PLoS Comput Biol*. 2011;7(3):e1002024.
21. Iranzo J, Puigbo P, Lobkovsky AE, Wolf YI, Koonin EV. Inevitability of genetic parasites. *Genome Biol Evol*. 2016;8(9):2856–69.
22. Raoult D, Forterre P. Redefining viruses: lessons from Mimivirus. *Nat Rev Microbiol*. 2008;6(4):315–9.
23. Koonin EV. The two empires and three domains of life in the postgenomic age. *Nat Educ*. 2010;3(9):27.
24. Abergel C, Legendre M, Claverie JM. The rapidly expanding universe of giant viruses: Mimivirus, Pandoravirus, Pithovirus and Mollivirus. *FEMS Microbiol Rev*. 2015;39(6):779–96.
25. Yutin N, Wolf YI, Koonin EV. Origin of giant viruses from smaller DNA viruses not from a fourth domain of cellular life. *Virology*. 2014;466–467:38–52.
26. Koonin EV, Wolf YI, Katsnelson MI. Inevitability of the emergence and persistence of genetic parasites caused by evolutionary instability of parasite-free states. *Biol Direct*. 2017;12(1):31.
27. Haruna I, Spiegelman S. Autocatalytic synthesis of a viral RNA in vitro. *Science*. 1965;150(3698):884–6.
28. Spiegelman S, Haruna I, Pace NR, Mills DR, Bishop DH, Claybrook JR, Peterson R. Studies in the replication of viral RNA. *J Cell Physiol*. 1967;70(2): Suppl):35–64.
29. Mills DR, Kramer FR, Spiegelman S. Complete nucleotide sequence of a replicating RNA molecule. *Science*. 1973;180(4089):916–27.
30. Mills DR, Peterson RL, Spiegelman S. An extracellular Darwinian experiment with a self-duplicating nucleic acid molecule. *Proc Natl Acad Sci U S A*. 1967;58(1):217–24.
31. McKean KA, Yourth CP, Lazzaro BP, Clark AG. The evolutionary costs of immunological maintenance and deployment. *BMC Evol Biol*. 2008;8:76.
32. Cotter SC, Topham E, Price AJ, Kilner RM. Fitness costs associated with mounting a social immune response. *Ecol Lett*. 2010;13(9):1114–23.
33. Parker BJ, Barribeau SM, Laughton AM, de Roode JC, Gerardo NM. Non-immunological defense in an evolutionary framework. *Trends Ecol Evol*. 2011;26(5):242–8.
34. Weinberger AD, Wolf YI, Lobkovsky AE, Gilmore MS, Koonin EV. Viral diversity threshold for adaptive immunity in prokaryotes. *MBio*. 2012;3(6): e00456–12.
35. Czarán T, Konnyu B, Szathmari E. Metabolically coupled replicator systems: overview of an RNA-world model concept of prebiotic evolution on mineral surfaces. *J Theor Biol*. 2015;381:39–54.
36. Szilagyi A, Zachar I, Scheuring I, Kun A, Konnyu B, Czarán T. Ecology and evolution in the RNA world dynamics and stability of prebiotic replicator systems. *Life (Basel)*. 2017;7(4):E48.
37. Gilbert W. The RNA world. *Nature*. 1986;319:618.
38. Robertson MP, Joyce GF. The origins of the RNA world. *Cold Spring Harb Perspect Biol*. 2012;4(5):a003608.
39. Higgs PG. Chemical evolution and the evolutionary definition of life. *J Mol Evol*. 2017;84(5–6):225–35.
40. Turner PE, Chao L. Prisoner's dilemma in an RNA virus. *Nature*. 1999; 398(6726):441–3.
41. Bazykin AD. *Nonlinear Dynamics of Interacting Populations*, vol. 11. Singapore: World Scientific; 1998.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

